

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-338986

(43)Date of publication of application : 08.12.2000

(51)Int.Cl.

G10L 11/02

G10L 15/04

G06F 3/16

G10L 15/00

(21)Application number : 11-150004

(71)Applicant : CANON INC

(22)Date of filing : 28.05.1999

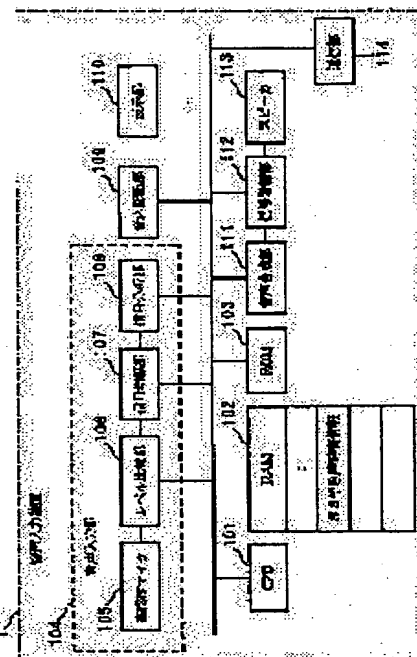
(72)Inventor : NISHIKAWA SHIGERU

(54) VOICE INPUT DEVICE, CONTROL METHOD THEREFOR AND STORAGE MEDIUM

(57)Abstract:

PROBLEM TO BE SOLVED: To wipe away a resistance sense of speaking to an apparatus thereby freely performing voice input, and also distinguish a voice vocalized to the apparatus from a voice vocalized to normal others.

SOLUTION: In this voice input device, a level comparison part 106 in a voice input part 1 compares an inputted voice signal with a threshold value for deciding whether the voice has whisper volume or not, and passes only a voice below the threshold. Thereby, a resistance sense to speaking to an apparatus is wiped away to freely perform voice input, while a voice vocalized to the apparatus can be distinguished from a voice vocalized to normal others.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

THIS PAGE BLANK (USPTO)

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-338986

(P 2 0 0 0 - 3 3 8 9 8 6 A)

(43) 公開日 平成12年12月8日(2000.12.8)

(51) Int. Cl. ⁷	識別記号	F I	テ-マコード (参考)		
G10L 11/02		G10L 3/00	513	A	5D015
15/04		G06F 3/16	320	E	9A001
G06F 3/16	320	G10L 3/00	513	B	
G10L 15/00			551	B	

審査請求 未請求 請求項の数 8 O L (全6頁)

(21) 出願番号 特願平11-150004

(22) 出願日 平成11年5月28日(1999.5.28)

(71) 出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72) 発明者 西川 成

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

(74) 代理人 100076428

弁理士 大塚 康德 (外2名)

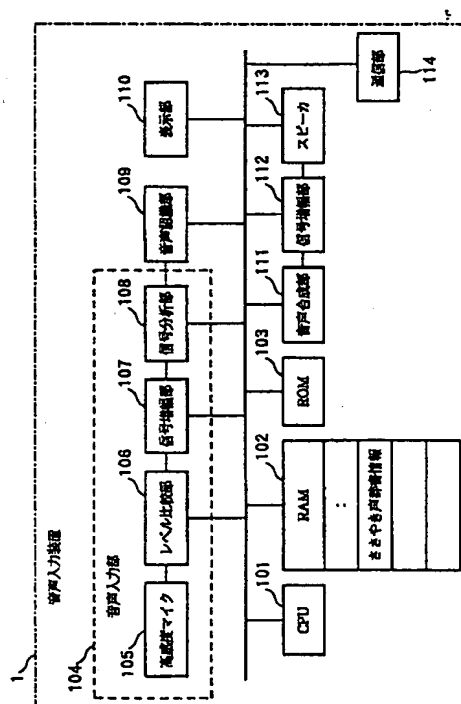
Fターム(参考) 5D015 CC13 CC14 CC18 KK03
9A001 BB02 BB04 HH17 HZ33

(54) 【発明の名称】 音声入力装置及びその制御方法及び記憶媒体

(57) 【要約】

【課題】 機器に話しかけることに対する抵抗感を払拭して気軽に音声入力を実施できるようにすると共に、通常の他者に対して発声した音声と機器に対して発声した音声を識別する。

【解決手段】 音声入力部1内のレベル比較部106は、入力された音声信号と、ささやき音声の音量であるか否かを判定するための閾値とを比較し、その閾値以下の音声についてのみ通過させる。



【特許請求の範囲】

【請求項1】 音声入力手段より入力された音声信号がささやき音声の音量であるか否かを判定するための閾値を有し、当該閾値を越える音声信号を非処理対象とし、前記閾値以下の音声信号について処理して出力する音声入力装置。

【請求項2】 更に、前記音声入力記号に対し入力信号が音声であるか否かを検出する音声信号検出手段を有し、当該音声信号検出手段によって音声であると判断された音声信号を処理対象とすることを特徴とする請求項1記載の音声入力装置。

【請求項3】 前記音声信号検出手段は、パワースペクトルとピッチ周波数により音声信号か否かを検出することを特徴とする請求項2記載の音声入力装置。

【請求項4】 更に、音声信号から音声を認識し、音声データを得る音声認識手段と、当該音声認識手段で得られた音声データに基づく情報を出力する出力手段とを備えることを特徴とする請求項第1項乃至第3項のいずれか1つに記載の音声入力装置。

【請求項5】 前記出力手段は、音声認識して得られた文字コードを、かな漢字変換処理に出力することを特徴とする請求項第4項に記載の音声入力装置。

【請求項6】 前記音声認識手段は、ささやき声で発声した場合の辞書パターンを有し、当該辞書パターンを用いて音声認識処理を行うことを特徴とする請求項4記載の音声入力装置。

【請求項7】 音声信号がささやき音声の音量であるか否かを判定するための閾値と、音声入力手段より入力された実音声信号とを比較し、当該閾値を越える音声信号を非処理対象とし、前記閾値以下の音声信号について処理して出力することを特徴とする音声入力装置の制御方法。

【請求項8】 コンピュータが読み込み実行することで、音声入力手段より入力された音声信号を処理する装置として機能する、プログラムコードを格納する記憶媒体であって、音声入力手段より入力された音声信号がささやき音声の音量であるか否かを判定するための閾値データを有し、当該閾値を越える音声信号を非処理対象とし、前記閾値以下の音声信号について処理して出力するプログラムコードを格納する記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は音声入力装置及びその制御方法及び記憶媒体に関するものである。

【0002】

【従来の技術】 音声入力により、機器を制御する装置としては、留守番メッセージや用件を音声で入力する留守番電話をはじめ、近年では音声認識機能付きのカーナビ等がある。また、キーボードに変わる文字入力手段とし

ての活用も行われている。しかしながら、このような音声入力装置を使用して、音声を入力する際に入力する音声を他者に聞かれたくない、もしくは他人に迷惑をかけたくないといった抵抗感を感じる状況がある。或いは、公衆の場で大きな声で話し辛い等もある。

【0003】 周囲の他者に音声情報を聴かれずに、また、周囲の雑音に影響されずに音声を入力することを實現するものとして、例えば、特開平5-265472号公報の音声情報処理装置、特開平5-114880号公報の携帯型移動無線端末といった提案がある。

【0004】 特開平5-265472号公報に開示された音声情報処理装置は、楕円体構造の音声反射器を用い、楕円体の一つの焦点に音声入出力のための送受話器（スピーカとマイクロホン）を、他の焦点に装置の操作者の頭部、特に耳と口がくる構造とすることにより、装置の操作者は、ハンドセットを持つことなく、かつ、後方に並んでいる他人に入出力情報を聴かれることもなく音声情報の入出力を行うことができるというものである。

【0005】 また、特開平5-114880号公報に開示された携帯型移動無線端末は、背景雑音の高いところでも使用者が大声を出さなくてすむ体積の小さい携帯型移動無線端末を提供する為に、予め話者の想定使用状況に応じて登録された複数の音声辞書を有する音声認識装置に話者音声を入力し、音声認識された単語出力を、話者が変換を希望する声調に応じて複数の音声辞書（I）が付随している音声合成装置に入力し、合成音声を出力するというものである。これにより、マイク入力から登録話者の音声として認識可能な音節のみを抽出していくことが可能であるので、体積の小さいマイクを使用していても背景雑音の高いところにおいて使用者が大声を出さずに通話可能になり、また、音声辞書に登録されていない音声は認識されないので、所有者以外の不正使用を実質的に禁止することができる。さらに音声認識と音声合成の間に同義文章変換の機能を付加したり、音声合成に際して声調変換を實行可能であるので通話相手に与える印象を制御でき、対話・交渉を有利に運べる効果がある。

【0006】

【発明が解決しようとする課題】 然るに、いずれの場合においても、周囲の雑音の影響を軽減して小さな音声でも確実に音声情報を機器に入力することを實現しているものの、機器に話し掛けることに対する抵抗感を払拭して気軽に音声入力を実施できるようにすること自体に注力しておらず、通常の他者に対して発声した音声と機器に対して発声した音声を識別するということは實現するものはなかった。

【0007】 本発明は、かかる問題に鑑みなされたものであり、機器に話しかけることに対する抵抗感を払拭して気軽に音声入力を実施できるようにすると共に、通常

の他者に対して発声した音声と機器に対して発声した音声を識別することができる音声入力装置およびその制御方法及び記憶媒体を提供することである。

【0008】

【課題を解決するための手段】かかる課題を解決するため、例えば本発明の音声入力装置は以下の構成を備える。すなわち。

【0009】音声入力手段より入力された音声信号がささやき音声の音量であるか否かを判定するための閾値を有し、当該閾値を越える音声信号を非処理対象とし、前記閾値以下の音声信号について処理して出力する。

【0010】また、本発明にかかる好適な実施態様に従えば、更に、前記音声入力記号に対し音声信号が音声であるか否かを検出する音声信号検出手段を有することが望ましい。これによって、小音量であっても、非音声の周囲の雑音を除外することができるようになる。

【0011】また、この音声信号検出手段は、パワースペクトルとピッチ周波数により音声信号が否かを検出することが望ましい。これによって、機械音と人間が発する音声とが識別できるようになる。

【0012】また、更に、音声信号から音声を認識し、音声データを得る音声認識手段と、当該音声認識手段で得られた音声データに基づく情報を出力する出力手段とを備えることが望ましい。

【0013】また、この場合の出力手段は、音声認識して得られた文字コードを、かな漢字変換処理に出力するようにすれば、これまでのキーボード等にとって代わる入力手段として機能させることが可能になる。

【0014】また、出力手段は、音声認識して得られた文字コードをコマンドとして変換処理に出力すれば、ボタンやスイッチの代わりに走査手段として機能させることも可能になる。

【0015】更に、音声認識手段は、ささやき声で発声した場合の辞書パターンを有し、当該辞書パターンを用いて音声認識処理を行うことが望ましい。この結果、ささやき声専用の辞書を用いることにより、ささやき声のみの認識精度を高めることができる。

【0016】

【発明の実施の形態】以下、添付図面に従って本発明にかかる実施形態を詳細に説明する。

【0017】＜音声入力装置の構成＞図1に本実施形態の音声入力装置のブロック図を示す。

【0018】図中、1は本実施形態の音声入力装置であり、以下の構成を有する。

【0019】101は、後述するROM103のプログラムに基づいて全体の制御を司るCPUである。102は、ささやき声辞書情報その他各種制御データ及びユーザデータを蓄積するRAM、103は、音声入力装置1の制御プログラムを格納するROMである。

【0020】104は、音声入力部であり、小音量の音

声でも感度よく音響電気変換を行う高感度マイク105、入力信号を所定の下限レベルと比較すると共に、所定の上限レベルと比較するレベル比較部106、信号を増幅する信号増幅部107、入力信号からピッチ周波数やパワースペクトルを算出し、音声信号の検出を行う信号分析部108で構成される。

【0021】109はCPU101の制御に基づきRAM102内のささやき声辞書情報を参照して音声認識を行う音声認識部、110は音声認識の結果等を操作者に通知する為の表示部である。111は音声認識の結果等を操作者に通知する為の合成音声を生成する音声合成部、112は音声合成部の出力合成音声を適切な信号レベルに増幅する信号増幅部、113は音声認識の結果等を操作者に通知する為の合成音声を出力するスピーカ（イヤホン）である。また、114は通信部であり、外部装置（例えばパーソナルコンピュータ等）との無線通信を行うものである。外部機器側にも通信部114と通信するためのデバイスが接続され、これを介して文章の入力を行う。

【0022】図2に、実施形態における音声入力装置の一形態である使用状況を示す。図示の如く、操作者には、マイク150及びスピーカ113が装着される形態を成している。

【0023】次に、図3のフローチャートに従って、本実施形態の音声入力動作を説明する。

【0024】本実施形態に於いては、予め、RAM102内のささやき声辞書情報を操作者のささやき声に適応させる為にトレーニングを行ってあるものとする。

【0025】先ず、ステップS101で、高感度マイク105からの信号入力を検出する為に、レベル比較部106に於いて所定の下限レベルと比較する。ステップS101に於ける比較の結果所定の下限レベル未満の場合はステップS101に戻り、所定の下限レベル以上の場合はステップS102に進む。

【0026】ステップS102では、小音量の音声信号のみ取り込む為に、レベル比較部106に於いて所定の上限レベルと比較する。この比較の結果、所定の上限レベルより大きい場合はステップS101に戻り、所定の上限レベル以下の場合はステップS103に進む。

【0027】この結果、通常の対話での音量以下の音声で、且つ、周囲の雑音を除くための音量以上の音声のみを通過させることができる。

【0028】ステップS103に進むと、信号増幅部108に於いて入力信号を適正な入力レベルに増幅する。そして、ステップS104において、信号分析部109に於いて入力信号の分析を行う。例えば、ピッチ周波数やパワースペクトラムを算出する。次いで、ステップS105において、ステップS104の分析結果より音声信号と判定できない場合は、ステップS101に戻り、音声信号と判定できる場合はステップS106に進む。

【0029】ステップS106では、音声認識部112に於いてRAM102内のささやき声辞書情報を参照して入力信号の音声認識を行う。そして、ステップS107で、ステップS106の音声認識結果に基づき、操作者に対する音声入力のリアクションを通知する。例えば、音声合成部112より音声入力に応じて求められる音声情報を出力したり、認識が正常に済んだことを知らせるための所定の音をスピーカ（イヤホン）で再生させたり、或いは、表示部111にリアクションとしての情報を表示する。

【0030】そして、ステップS108において、通信部114を介して外部装置に認識結果（文字コード群）を通知する。

【0031】本実施形態の音声入力装置は、通常のキーボードに代わる入力手段として機能させることができる。また、スイッチやボタンに代わるコマンド入力手段としても機能させることもできる。

【0032】外部装置が例えばパーソナルコンピュータである場合、通信部114からの情報（かなを示す文字コード群）を受信し、それをかな漢字変換プログラム（FEP等と呼ばれている）或いはコマンド解釈プログラムに引き渡すデバイスドライバ（プログラムの一種）として動作させておけばよい。

【0033】また、本装置が携帯できる装置である場合には、通信部114としては無線による通信が望ましいが、物理的なケーブルで接続する形態としてもよいのは勿論である。

【0034】また、実施形態では、音声認識し、それをキーボードの代わりに文字を入力する、或いはスイッチの代わりにコマンドを入力する例を説明した。しかしながら、本発明は幅広い適用が考えられ、これによって本発明が限定されるものではない。

【0035】更に、昨今のパーソナルコンピュータの情報処理装置には、標準で音源ボードが装着されており、そこにはマイク接続端子が備わっている。このマイク接続端子にマイクロホンを接続し、その音源ボードとパーソナルコンピュータ内のCPU及び図3の処理を実現するプログラムでもって実現するようにしてもよい。または、好感度のマイクを接続する専用の拡張ボードをパーソナルコンピュータに装着して実現してもよい。

【0036】この場合のプログラムはデバイスドライバ、もしくはアプリケーションプログラムの一部として動作することになるが、図3と異なるのは、ステップS108における出力対象が異なるのみである。

【0037】以上、説明したように、本実施形態に於いては、音声入力装置に於いて、所定音量以下の音声信号のみ入力する構成と、前記入力信号に対し音声信号であることを検出する音声信号検出部と、前記音声信号検出部にパワースペクトルとピッチ周波数により音声信号を検出する構成と、前記入力信号に対して音声認識処理を

行う音声認識部と、前記音声認識部にささやき声で発声した場合の辞書パターンを有し、当該辞書パターンを用いて音声認識処理を行う構成とを設けることにより、機器に話し掛けることに対する抵抗感を払拭して気軽に音声入力を実施できるようにすると共に、通常の他者に対して発声した音声と機器に対して発声した音声を識別することができる音声入力装置および方法を提供するという大きな効果がある。

【0038】本実施形態に於いては、音声入力装置の高感度マイクと本体の形態は図2に示した概観と異なっており、例えば、マイクをクリップ型にして襟にはさんで使用するような形態等他の形態であっても構わないということは言うまでもない。

【0039】また、図3におけるステップS101、102で比較される閾値は操作者によって適宜調整できるようにしてある（不図示のボリュームツマミを操作する）。これにより、ささやき声の個人差、或いは周囲の雑音の状況の問題を吸収できるようになる。

【0040】また、本発明は、複数の機器から構成されるシステムに適用しても、一つの機器からなる装置（例えばマイク内蔵型情報処理装置）に適用してもよい。

【0041】また、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体（または記録媒体）を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム(OS)などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0042】さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0043】

【発明の効果】以上説明したように本発明によれば、機器に話し掛けることに対する抵抗感を払拭して気軽に音声入力を実施できるようにすると共に、通常の他者対

して発声した音声と機器に対して発声した音声を識別することができるようになる。

【図面の簡単な説明】

【図 1】 音声入力装置ブロック図である。

【図 2】 音声入力装置概観図である。

【図 3】 音声入力動作フローチャートである。

【符号の説明】

1 音声入力装置

101 CPU

102 RAM

103 ROM

104 音声入力部

105 高感度マイク

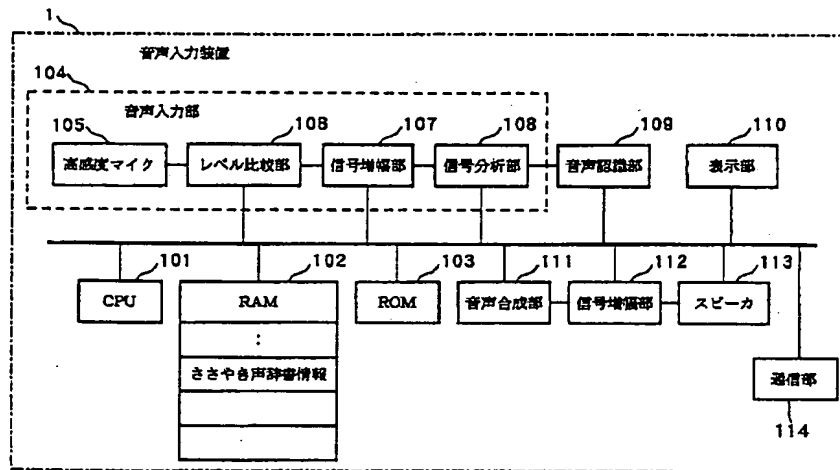
106 レベル比較部

107 信号増幅部

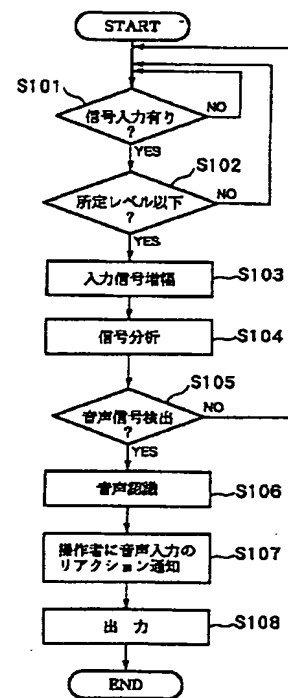
108 信号分析部

109 音声認識部

【図 1】



【図 3】



【図2】

